Isotope

USB HID Emulation for Embedded Devices

Benjamin Matthew Pannell

**STUDY LEADER** Professor Thomas Niesler

**DATE** August 2014

**Previous Works**

Pi Your Command

*Automatic Speech Recognition using the Raspberry Pi*

Christian Truter

# Summaries

Isotope is a project which addresses the need for an easy to use, low cost, USB HID emulation framework for use on embedded devices as an interface between these devices and any personal computer.

Applications include, but are not limited to, voice control of personal computers to aid performance and usability while ensuring universal compatibility. It is also possible that Isotope may be used to rapidly develop low cost simulator controls, remote control devices and administration tools.

Isotope has been designed to make use of the low cost ATmega32u4 chip which is readily available and can be sourced in small volumes for easy prototyping. Hardware integration has been kept as simple as possible, and maximum flexibility with respect to the host device has been sought to allow future expansion.

# Table of Contents

[Summaries 2](#_Toc395540662)

[Table of Contents 3](#_Toc395540663)

[Figures 4](#_Toc395540664)

[Tables 4](#_Toc395540665)

[Symbols 5](#_Toc395540666)

[Introduction 6](#_Toc395540667)

[Pre-Design Investigation 8](#_Toc395540668)

[Device Design 10](#_Toc395540669)

[Component Information 12](#_Toc395540670)

[References 13](#_Toc395540671)

[Appendix A: Project Planning Schedule 14](#_Toc395540672)

[Appendix B: Project Specifications 15](#_Toc395540673)

[Appendix C: Outcomes Compliance 16](#_Toc395540674)

[Appendix D: Circuit Diagram 17](#_Toc395540675)

[Appendix E: Performance Benchmarks 18](#_Toc395540676)

[Appendix F: Communications Protocol 19](#_Toc395540677)

[Appendix G: C-Library API 24](#_Toc395540678)

[Appendix H: Source Code 25](#_Toc395540679)

# Figures

[Figure 1 Arduino Nano. Image courtesy of Arduino SA 9](#_Toc395540680)

[Figure 2 PJRC Teensy 2.0. Image courtesy of PJRC 9](#_Toc395540681)

[Figure 3 TXB0104 Bi-Directional Level Shifter. Image courtesy of Adafruit Industries 10](#_Toc395540682)

[Figure 4 Adafruit Prototyping Pi Plate Kit for Raspberry Pi. Image courtesy of Adafruit Industries 11](#_Toc395540683)

[Figure 5 Raspberry Pi Model B. Image courtesy of Adafruit Industries 11](#_Toc395540684)

[Figure 6 Circuit Diagram 17](#_Toc395540685)

# Tables

[Table 1 Component List 12](#_Toc395540686)

[Table 2 Component Websites 12](#_Toc395540687)

[Table 3 UART Protocol Packet Structure 20](#_Toc395540688)

[Table 4 UART Protocol OP Codes 20](#_Toc395540689)

# 

# Symbols

**Adafruit Industries** – An online electronics hobby store focusing on embedded device development and wearable computing.

**Application Programming Interface** - A set of methods, often provided in the form of a library, enabling 3rd party applications to make use of another application, device or service.

**Application Specific Integrated Circuit** – A chip designed to serve a specific purpose, in contrast to programmable devices which may be modified to suit a number of purposes.

**Bipolar Junction Transistor** – A form of switch activated by the flow of current through a base node. Similar in purpose to a FET.

**Central Processing Unit** – The primary component of a programmable chip, responsible for performing most logical operations as a result of the instructions it receives.

**Field Effect Transistor** – A form of switch activated by the voltage difference over its gate node. Similar in purpose to a BJT.

**Future Technology Devices International** – A hardware design company responsible for a number of USB interface chips used in a wide range of devices.

**Human Interface Device** – A subset of the USB protocol which makes allowance for the use of devices through which humans can interact with their computers, removing the need for dedicated drivers.

**Printed Circuit Board**– Circuit boards on which copper “tracks” are printed to simplify the construction of complex circuitry.

**Random Access Memory** – High speed memory used to store data which is currently in use on the system, including the instructions of running processes and the information they are interacting with.

**Universal Asynchronous Receiver Transmitter** – A standardised device responsible for serial communication between two devices over a two-wire connection.

**Universal Serial Bus**– A serial communication protocol, and accompanying hardware specification, which dictates the way in which a number of different device types may be connected to computers.

API *See* Application Programming Interface

ASIC *See* Application Specific Integrated Circuit

BJT *See* Bipolar Junction Transistor

CPU *See* Central Processing Unit

FET *See* Field Effect Transistor

FTDI *See* Future Technology Devices International

HID *See* Human Interface Device

PCB *See* Printed Circuit Board

RAM *See* Random Access Memory

UART *See* Universal Asynchronous Reciever Transmitter

USB *See* Universal Serial Bus

# Introduction

Modern voice recognition systems commonly fall into two primary categories, cloud based and native. Examples of cloud based speech recognition engines include Google’s Voice Search, Apple’s Siri virtual assistant and more recently, Microsoft’s Cortana. The native implementations are best represented by Nuance’s Dragon series of products and Microsoft’s proprietary Speech API (MSSAPI).

Native solutions are generally built on learning Hidden Markov Models which adapt to the speaker and can often achieve high accuracy levels once trained and paired with a high quality microphone. In most cases these systems are designed to assist people who would otherwise be required to perform a lot of typing, or the disabled, and as a result their implementations are often tailored towards single users.

Conversely, cloud solutions are vastly more complex and generally designed to be able to achieve good accuracy rates with little or no speaker specific adaptations. As a result they are often built using a combination of advanced neural networks and HMMs to help improve feature detection across very large datasets, such as those which Google acquired through their 1-800-GOOG-411 service [1]. There are a few restrictions to these cloud based services though, often determined by their target applications, primarily limitations on the maximum length of a dictated statement and the inability to adapt to a user’s pronunciation. In practice however, these limitations are minor in nature.

One of the major issues faced with both systems is that they rely on software on the target device to record, pre-process, recognise where necessary, and finally output the result – leading to platform restrictions which are often difficult to overcome. Another is the possibility of piracy, as this software is often extremely expensive with a low number of users it poses a major threat to the producer’s revenue stream.

This project hopes to enable the development of a device which addresses both issues in an elegant and universally compatible manner, namely through the emulation of a user’s input devices. This will allow speech recognition to be performed either on a hardware processor attached to the user’s computer or on the cloud with this device as a proxy. As the solution is hardware based, piracy will be impossible and the revenue stream of the producer will be more secure.

In addition to this, the ability to easily move the device between any computer, has the advantage of allowing learning algorithms to be applied – either on the device itself or using the device as an identifier - improving recognition rates for the device’s user.

The goal of this project is to develop an interface which can be used by embedded devices to emulate a user’s input devices – such as a keyboard or mouse – without the installation of custom drivers or software on the target machine. In light of this requirement, this project has taken the form of a USB HID emulation chip which is controllable over a simple serial protocol over an UART, allowing it to be used on almost any embedded device platform with minimal, or in some cases no, hardware alterations.

As an adjunct to this, a series of communications protocols and libraries will be developed to make the use of the emulation hardware as straightforward as possible from a variety of different programming languages. To demonstrate this, a simple speech recognition engine will be implemented to allow basic commands to be given to the device and executed.

# Pre-Design Investigation

Prior to beginning design of the system it was important to investigate the possible approaches and determine which of them best suited the task at hand. During this phase a number of possible solutions were investigated, their advantages and disadvantages compared and finally a decision was made on the best option for this project.

## Custom ASIC Design

The first option to be considered was the design of a custom ASIC for the purpose of USB emulation, implementing its own interface over either serial UART or i2c. Doing so would have allowed the manufacture of extremely small, energy efficient and cheap emulation chips and would have proven an ideal solution for mass production due to the potential cost and size savings involved. On the other hand, by virtue of the solution being entirely implemented in hardware, the design and testing phase would have been prolonged while simultaneously reducing the flexibility with which additional features could be added.

To summarise, a custom ASIC providing USB HID emulation would suit a large scale project, however for prototyping purposes it would be present challenges, difficult to acquire and force developers to spend a large amount of time implementing support for its low level interface. In particular, the design and implementation of the low level interface would be time consuming, with later modifications posing the same issue.

## Programmable USB Slave Device

There are a number of USB interface chips available on the market, for example – the well-known FT232R [2] by FTDI - but also more complex programmable devices. One such example is the Vinculum-II [3] which includes a built in 16-bit CPU and programmable code block, allowing you to easily modify it to suit any number of applications.

The primary reason for avoiding the use of these devices was the difficulty of obtaining a chip which made use of common slave-type (B class) connectors in a pre-packaged form. This would require the purchase of individual chips and – due to their form factors – the surface mounting on custom PCBs, restricting the ability to construct the device easily and potentially raising costs above acceptable levels for small runs.

## Microprocessor with USB Interface

Another option was to make use of a microprocessor which included a built in USB interface, reprogramming its firmware to allow USB HID emulation to take place and repurposing one of its IO channels for inter-device communication. One particularly promising candidate was the ATmega32u4 [4] which included its own full speed USB controller which was fully programmable. Other advantages included the fact that it was the basis of the Arduino Nano [5] and as a result was commonly available in an easy to use package.

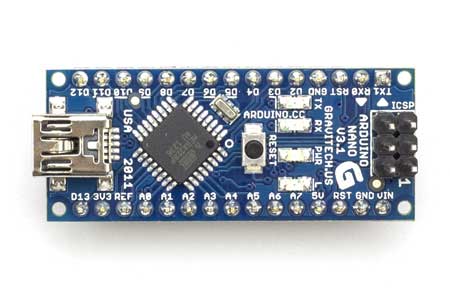


Figure Arduino Nano. Image courtesy of Arduino SA [5]

In addition to this, the ATmega32u4 is designed to act as a controller for a number of USB peripherals. This proven functionality improved confidence in the chip’s ability to fulfil the requirements at hand, while its small size and relatively low cost would ensure low costs in production. In addition to this, it would be easily possible to transition to either the ATmega16u2 or ATmega8u2 in production without major code changes – allowing costs to be reduced further.

## Conclusion

After analysing three possible options, it was decided that the ATmega32u4 would provide the best prototyping platform for the project with the best prospects for future expansion, while remaining accessible and low cost.

In selecting the ideal prototyping platform, it was noted that the Arduino Nano [5] provided adequate functionality. However the PJRC Teensy 2.0 [6] provided a more rounded set of features at a lower cost, while remaining equally accessible and using the same tooling. In addition to this, the Teensy 2.0’s size – roughly one third smaller than the Arduino Nano – meant the final prototype would be smaller and more portable. As a result the decision was made to acquire the Teensy 2.0 as the prototyping board of choice.

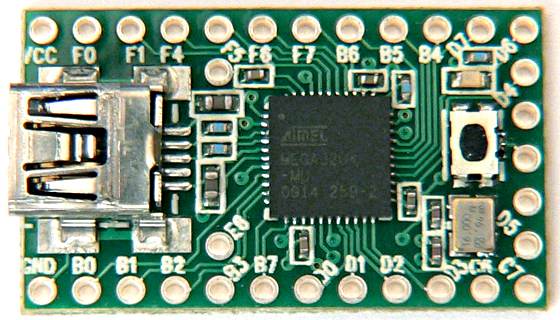


Figure PJRC Teensy 2.0. Image courtesy of PJRC [6]

# Demonstration Device Selection

In order to demonstrate the emulation device, it was necessary to acquire a master device similar to that which would be used in production or application prototyping environments. Due to Professor Thomas Niesler’s experience with speech recognition, and interest in making use of the device in this field, it was decided that the demonstration would consist of a self-contained speech recogniser – requiring that the master device be capable of running the software required.

Following on the work performed by Christian Truter in implementing Automatic Speech Recognition on the Raspberry Pi in his thesis project Pi Your Command [7] it was decided that a Raspberry Pi class device would suffice as a low cost demonstration platform for a number of reasons discussed below.

There are a number of devices of similar performance occupying a similar price point to the Raspberry Pi [8] including the BeagleBone Black [9] and, although somewhat more expensive, the Intel Galileo Development Board [10]. More recently, Raspberry PI class boards such as the Banana Pi [11] and HummingBoard [12] have also begun to appear which offer a higher performance solution at a corresponding price point.

Of these devices, one of the primary concerns within the scope of this project was the cost effectiveness of the prototyping platform. While options such as the HummingBoard [12] are significantly more powerful than the Raspberry Pi [13], this additional processing power would not make any significant contribution to this project. In addition to this, the popularity of the Raspberry Pi has served to ensure that devices such as the HummingBoard and Banana Pi adopt the same physical layouts. This allows the design of this project’s module to be specifically tailored towards a specific form factor, allowing drop-in replacement of the master board.

Availability, both of the physical hardware as well as support, examples and documentation also played an important role in the decision making process - a field in which the Raspberry Pi’s popularity is again a factor.

# Device Design

There were a number of device aspects which need to be taken into account when undertaking design of the board and its associated components. These would dictate the ways in which components were connected and have an effect on the communication protocols used to allow the Raspberry Pi [7] to communicate with the Teensy 2.0 [6].

One of the initial design considerations was based on the fact that the Raspberry Pi [7] operated at a core voltage of 3.3V while the Teensy 2.0 [6] – for lack of a voltage converter – operated at 5.0V supplied via its USB port. As a result of these different voltage levels, it would be dangerous to connect the Raspberry Pi [7] and Teensy 2.0 [6] directly to one another.

For this reason it would be necessary to include a level translator in the design to allow the Raspberry Pi [7] to communicate safely with the Teensy 2.0 [6]. Initially the use of a voltage divider and basic BJT or FET booster were considered [8], however upon further inspection it became clear that at very high switching rates the voltage divider may become unsuitable due to the output pin capacitances. The BJT or FET booster and voltage divider would also need to be tailored to the voltages of each device, reducing the ease with which the system could be adapted to new host hardware.

[9]

[10]

Seeking an alternative solution, the Texas Instruments TXB0104 4-channel bi-directional level translator [9] was selected based on its ability to handle a wide range of voltages from 1.2V to 3.6V on the low side and 1.65V to 5.5V on the high side [10], with automatic direction detection on each channel. This, combined with the exceptionally high throughput (100Mbps [10]) meant that it would be possible to allow easy migration of the final device between different host devices with minimal, if any, modifications as well as allowing future extensions to the device’s capabilities as necessary.

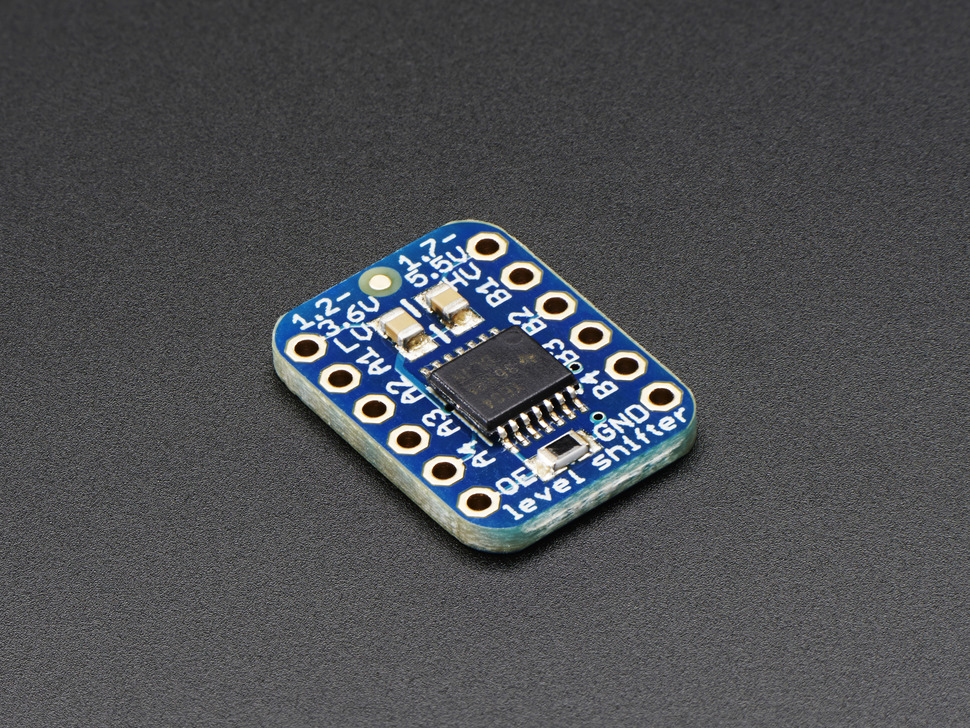


Figure TXB0104 Bi-Directional Level Shifter. Image courtesy of Adafruit Industries [11]

Finally, to assist with prototyping on the Raspberry Pi [7] it was decided that a platform specific prototyping board, in the form of the Adafruit Prototyping Pi Plate Kit [11], would be used to allow easy attachment to the Raspberry Pi [7] and provide a stable platform on which to mount the Teensy 2.0 [6] and TXB0104 [9].

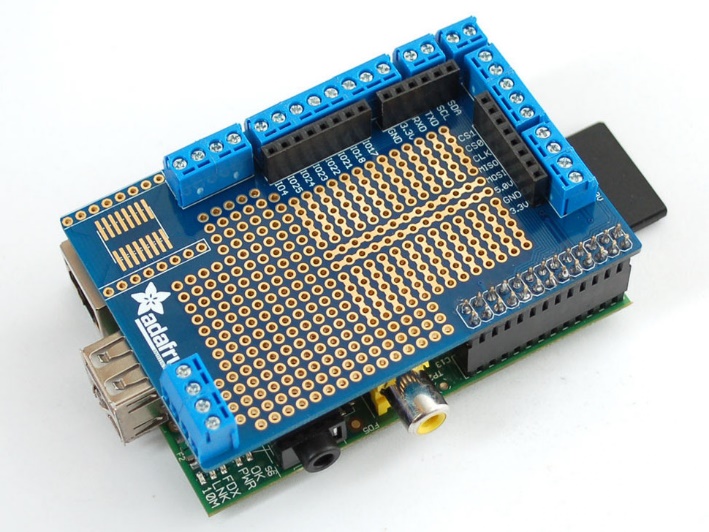


Figure Adafruit Prototyping Pi Plate Kit for Raspberry Pi. Image courtesy of Adafruit Industries [12]

## Electrical Interface

The primary hardware design challenge is the electrical interface between the master device (Raspberry Pi in this case) and the ATmega32u4. The design requires that a one-way command channel be established between these devices in a reliable and safe manner, requiring minimal setup on behalf of the user.

### Hardware Protocol Selection

There was also a strong incentive to make use of common standardized interface protocols like i­2c, SPI or UART in order to allow the design to easily be adapted to alternative master devices. When considering these protocols it was important to take a number of concerns into account, namely available bandwidth, voltage levels and switching frequencies.

In all cases, the differing core voltages utilized by the Raspberry Pi and ATmega32u4 (3.3V and 5V respectively) would require some form of voltage level switching circuit to be used.

When considering i2c and SPI, one of the immediately apparent issues was the synchronous nature of the link – requiring the master device to continuously poll the client for new data – which would complicate attempts to extend the interface to operate in a duplex manner. The second significant hurdle was the switching rates used by these interfaces, often in the order of several MHz, this would make the task of level switching vastly more complex as a non-resistor based solution would need to be sought to avoid a high RC time constant or current drain.

In addition to these issues, i2c requires a somewhat more complex pull-up arrangement which makes level translation more challenging for automated circuits, requiring specialist chips to ensure that it operates correctly. SPI, while not suffering from this issue, has the problem of a minimum of 3-wires, increasing the number of voltage level translators necessary and therefore the complexity of the design.

As a result, it was decided that UART offered a good compromise between available bandwidth (115200 baud would offer enough bandwidth to adequately convey commands with a binary protocol), simplicity (with only one wire for simplex, and two for duplex communication) and availability (with most integrated devices sporting at least one UART).

### Voltage Level Translation

Due to different core operating voltages on the Raspberry Pi and ATmega32u4 (Teensy 2.0) it is necessary to perform voltage level translation on the interface lines so as to prevent damage to the master or slave devices.

Two solutions were considered to this problem, the first being a custom BJT/resistor ladder circuit which would raise and lower (respectively) the voltage levels on the transmit and receive lines, and the second being a Texas Instruments TXB0104 [17] level translation chip which was designed for this purpose.

The advantage of a BJT/resistor ladder configuration was a significant reduction in project cost, by approximately 90% for the level switcher circuit, however it would impose restrictions on the flexibility of the design. Specifically, the BJT’s configuration would be tied to the master and slave voltage levels, and it would require replacement if either of these operating voltages changed in future. The same issue applied to the resistor ladder as a step-down converter, as well as possible transient interference with higher UART baud rates if resistor selections were poor – a mistake someone unfamiliar with the project could easily make when attempting their own implementation.

The TXB0104 [17] conversely offers an increase in flexibility, automatically translating voltage levels between 1.2V and 5.5V without any circuit modifications, and with automatic direction detection. This would significantly simplify circuit design and allow the project to be easily implemented by even inexperienced hobbyists at the result of a higher unit cost.

## Safety Precautions

In the interest of safety, it has been decided that the Raspberry Pi will not have its 5V connection coupled to the Teensy 2.0’s 5V connection, which would potentially allow the Raspberry Pi to operate without the need of an external power source. This decision was made in the interest of caution as the Raspberry Pi is capable of drawing far more than the USB 2.0 standard 500mA of current which, if drawn through the ATmega32u4, has the potential to damage the Teensy 2.0.

The result is that it is necessary for the Raspberry Pi to be connected to its own power source, independently of the Teensy 2.0.

## Raspberry Pi Selection

At the time of selection, there were two models of Raspberry Pi available. The Model A and the Model B. The Model B was selected due to the presence of an Ethernet Port, as well as the additional 256MB of RAM, with the intention of running a basic voice recognition framework on the device.

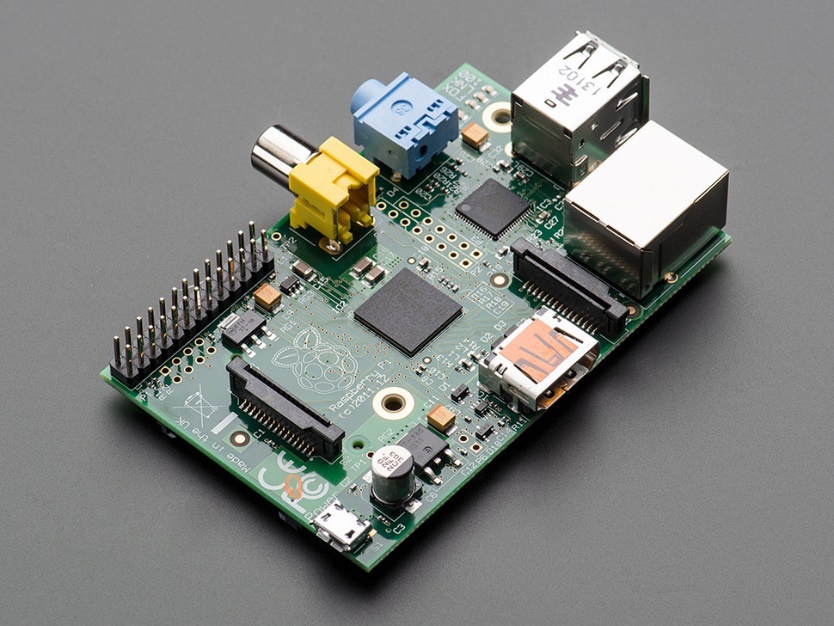


Figure Raspberry Pi Model B. Image courtesy of Adafruit Industries [15]

# Software Design

## Slave Device

The slave device, an ATmega32u4, is responsible for the conversion of emulation commands into USB HID packets which are conveyed to a computer over a USB connection. The PJRC Teensy 2.0 [6] selected for this purpose is bundled with a basic USB Mouse and Keyboard emulation library which is capable of emulating simple key presses and mouse movements.

Figure 6 Slave Device Command Process

The slave device’s software is therefore responsible for the parsing of a received packet, delegation of the parsed packet to the relevant logic function and finally the calling of the emulation functions required to fulfil the command.

In order to maximize throughput, minimize processing time and boost flexibility it was decided that a binary protocol represented the best approach to command structuring. The design of this protocol, its structure and examples are available in Appendix F: Communications Protocol.

## Master Device

The master device software consists of libraries responsible for wrapping of emulation commands in their relevant protocol level representations and transmission of these to the slave device. These libraries are available in a number of languages and, at a minimum, provide a means to submit mouse and keyboard emulation commands to the slave device for emulation.

Figure 7 Master Device Library Process

At a base level, these libraries provide functions for emulating mouse key presses, movement and scrolling through a function like isotope\_mouse(buttons, deltaX, deltaY, deltaScroll) and the emulation of keyboard input through a function like isotope\_keyboard(modifiers, [] keys). These functions are intended to be used by higher-level wrappers which extend their behaviour through the use of state machines, key maps and other application specific logic to provide more advanced functionality.

This approach allows complex functionality to be implemented at a library level while abstracting the details of the emulation away from the client – vastly simplifying the task of maintenance and future extension, as well as allowing differentiation within client libraries and customization of their output.

For example, it would be possible for a custom library to provide a transcription service through which text passed to it in string form would be emulated, effectively transcribing the text onto a target computer.

# Component Information

All components were sourced from Adafruit Industries, through their online store, and shipped to South Africa using United Parcel Service Worldwide Expedited shipping.

Table Component List

|  |  |  |
| --- | --- | --- |
| **Component Name** | **Price** | **Quantity** |
| Raspberry Pi Model B 512MB RAM [15] | $39.95 | 1 |
| Teensy 2.0 – ATmega32u4 [16] | $15.95 | 1 |
| Adafruit Prototyping Pi Plate Kit for Raspberry Pi [14] | $15.95 | 1 |
| TXB0104 Bi-Directional Level Shifter [13] | $3.50 | 1 |
| Extra-long break-away 0.1” 16-pin strip male header (5-pieces) | $3.00 | 1 |
| USB cable – 8” A to Mini B Charging and Micro B Data | $3.95 | 1 |

Table Component Websites

|  |  |
| --- | --- |
| **Component Name** | **Website** |
| Raspberry Pi Model B | http://www.raspberrypi.org/product/model-b/ |
| Teensy 2.0 | https://www.pjrc.com/teensy/ |
| TXB0104 | http://www.ti.com/product/txb0104 |
| Prototyping Plate | https://www.adafruit.com/products/801 |

# Future Expansion

This initial revision of the device represents the most basic level of functionality required to allow USB input device emulation in an embedded device for speech recognition. There are a number of additional features which could be added at a later stage including a built in display and/or menu to allow customization of the device and/or the display of contextual information like processed commands.

It would also be possible to integrate the text-to-keypress conversion functionality on-device, with a simple protocol extension allowing for transition between different keyboard layouts, thus removing the need to include this in implementation libraries and therefore offloading the primary device.

# References

|  |  |
| --- | --- |
| [1] | J. C. Perez, “Google wants your phonemes: InfoWorld,” InfoWorld, 23 October 2007. [Online]. Available: http://www.infoworld.com/t/data-management/google-wants-your-phonemes-539. [Accessed 26 07 2014]. |
| [2] | FTDI, “FT232R,” Future Technology Devices International Ltd., [Online]. Available: http://www.ftdichip.com/Products/ICs/FT232R.htm. [Accessed 07 August 2014]. |
| [3] | FTDI, “Vinculum II,” Future Technology Devices International Ltd., [Online]. Available: http://www.ftdichip.com/Products/ICs/VNC2.htm. [Accessed 07 August 2014]. |
| [4] | Atmel Corporation, “ATmega32u4,” Atmel, [Online]. Available: http://www.atmel.com/devices/atmega32u4.aspx. [Accessed 11 August 2014]. |
| [5] | Arduino, “Arduino Nano,” Arduino, [Online]. Available: http://arduino.cc/en/Main/arduinoBoardNano. [Accessed 2014 August 11]. |
| [6] | PJRC, “Teensy 2.0,” PJRC, [Online]. Available: https://www.pjrc.com/store/teensy.html. [Accessed 2014 August 11]. |
| [7] | C. Truter, “Pi Your Command,” Stellenbosch University, Stellenbosch, 2013. |
| [8] | Raspberry Pi Foundation, “Raspberry Pi,” Raspberry Pi Foundation, [Online]. Available: http://www.raspberrypi.org/. [Accessed 11 August 2014]. |
| [9] | IanH, “RPi GPIO Interface Circuits,” eLinux.org, 30 August 2013. [Online]. Available: http://elinux.org/RPi\_GPIO\_Interface\_Circuits. [Accessed 11 August 2014]. |
| [10] | Atmel, “ATmega32u4 Datasheet,” [Online]. Available: http://magicshifter.net/static/datasheets/atmega32u4.pdf. [Accessed 18 August 2014]. |
| [11] | Mosaic Industries, “Raspberry Pi GPIO Pin Electrical Specifications,” Broadcom, [Online]. Available: http://www.mosaic-industries.com/embedded-systems/microcontroller-projects/raspberry-pi/gpio-pin-electrical-specifications. [Accessed 18 August 2014]. |
| [12] | Texas Instruments, “TXB0104,” Texas Instruments, [Online]. Available: http://www.ti.com/product/txb0104. [Accessed 11 August 2014]. |
| [13] | Texas Instruments, “TXB0104 Technical Documentation,” May 2012. [Online]. Available: http://www.ti.com/lit/gpn/txb0104. [Accessed 11 August 2014]. |
| [14] | Adafruit Industries, “TXB0104 Bi-Directional Level Shifter,” Texas Instruments, [Online]. Available: https://www.adafruit.com/products/1875. [Accessed 18 August 2014]. |
| [15] | Adafruit, “Adafruit Prototyping Pi Plate Kit for Raspberry Pi,” Adafruit, [Online]. Available: https://www.adafruit.com/products/801. [Accessed 11 August 2014]. |
| [16] | Adafruit Industries, “Raspberry Pi Model B 512MB RAM,” Raspberry Pi Foundation, [Online]. Available: https://www.adafruit.com/product/998. [Accessed 18 August 2014]. |
| [17] | Adafruit Industries, “Teensy (ATmega32u4 USB dev board) 2.0,” PJRC, [Online]. Available: https://www.adafruit.com/products/199. [Accessed 18 August 2014]. |
| [18] | B. Goodwine, “RS-232 Serial Protocol,” 29 September 2002. [Online]. Available: http://controls.ame.nd.edu/microcontroller/main/node24.html. [Accessed 11 August 2014]. |

# Appendix A: Project Planning Schedule

1. Functional Specification
2. Existing Implementation Research
3. Initial Design
4. Feasibility and Limitation Study
5. Refinement of Design, Parts List
6. Ordering of Components
7. Initial Software Design
8. Assembly
9. Software Testing
10. Platform Expansion – new libraries, demos, documentation
11. Voice Recognition Implementation

# Appendix B: Project Specifications

## Non-Functional Specifications

### N-001 No Custom Drivers

Device must make use of the USB HID protocol for communication to ensure that it does not require the development and installation of custom drivers.

### N-002 Keyboard Emulation

Device must be able to emulate a keyboard, allowing the pressing and releasing of individual keys, as well as combinations of keys and/or modifiers like Ctrl, Alt or Shift.

### N-003 Mouse Emulation

Device must be able to emulate a mouse, supporting movement of the mouse cursor in the X and Y axis, pressing and releasing of the Left, Middle and Right mouse buttons and scrolling up or down by emulation of a scroll wheel.

### N-004 Raspberry Pi Integration

Device must be able to be used in conjunction with a Raspberry Pi Model B, with no modifications to the Raspberry Pi itself so as to minimize the risk of damage.

### N-005 High Compatibility

Device must make use of a commonly available communication interface available on most embedded systems to allow adaptation to different platforms in the future.

## Functional Specifications

### F-001 High Performance

The device and protocol should ensure that any performance limitations are dictated by the USB HID protocol rather than the implementation. Specifically, the inter-device communication bus should be able to transmit at least 1000 commands per second.

### F-002 Low Power Usage

The device should not require more power than can be supplied through the USB port of the system it is connected to – 500mA on USB2.0 and 1200mA on USB2.0 High-Power or USB3.0.

### F-003 Low Cost

The device should be cheap to prototype as well as having low cost components such that mass production profit margins may be maximized. Maximum cost of the emulation components for prototyping may not exceed R500 while production costs shall not exceed R100.

### F-004 Small Size

The device should require a minimum of PCB surface area such that production devices may be built to be extremely small and lightweight so as to be easily portable.

# Appendix C: Outcomes Compliance

The following ECSA outcomes have been satisfied in the listed sections.

# Appendix D: Circuit Diagram

The circuit diagram for the Isotope emulation device, as connected to the Raspberry Pi’s UART, is show below. Take note that at no point are the Raspberry Pi and Teensy 2.0’s power supplies coupled, ensuring that the Raspberry Pi cannot draw current through the Teensy 2.0 – as doing so has the possibility of damaging the Teensy 2.0’s circuitry.

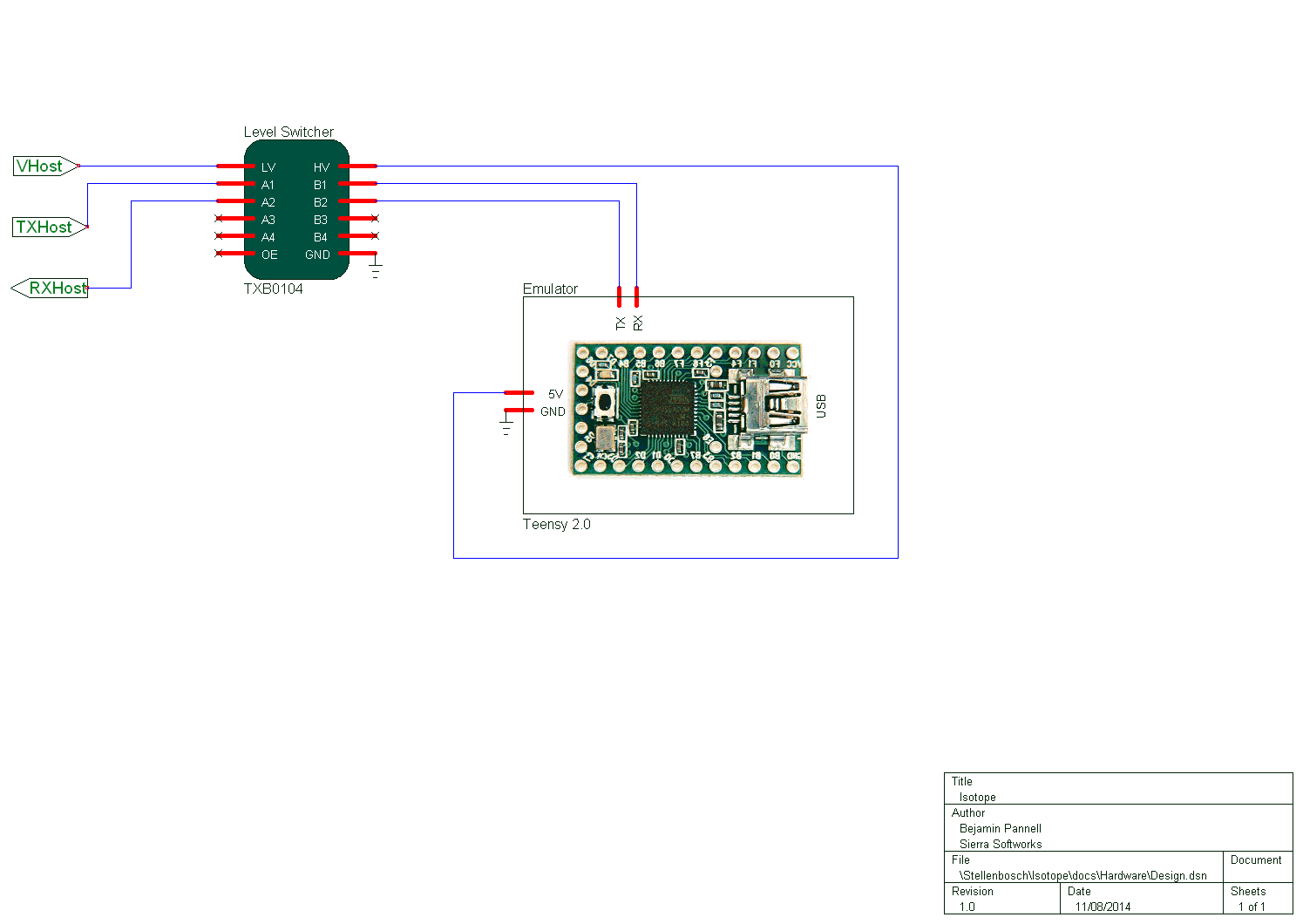


Figure Circuit Diagram

# Appendix E: Performance Benchmarks

Need to do performance testing to determine the maximum data throughput, as well as the best baud rate to operate the UART at to help prevent data-loss due to buffer overflows.

Due to the Teensy’s firmware using a circular buffer to store messages, buffer overflows do not pose a security issue.

## USB Protocol Limitations

The USB HID protocol specifies that for each HID device connected to the host, provision must be made to accept up to 1000 reports per second (1000Hz polling rate). As this is a protocol specific limitation, we will use it as the upper limit on our performance benchmarks (exceeding this limit should not be permitted).

## UART Limitations

The UARTs used on the Raspberry Pi and Teensy 2.0 are capable of operating at baud rates of up to 115200 baud, as we are using 8-bit packets with even parity we are in fact sending 11-bits per packet. [12] From this it is possible to deduce that our maximum packet throughput is equal to .

## Isotope Protocol Limitations

Due to the flexible design of the Isotope protocol, in which packet lengths can vary from 1-byte to 32-bytes depending on their context, it is important to analyse performance given different usage scenarios.

The most common usage scenario is text input, in which one key is pressed at a time (with optional modifier keys) and then released. Optimizations to the way in which key presses are sent (only sending the key release command if the key remains the same, or at the end of the text) can help improve throughput, however for the sake of our calculations we will assume that each key is pressed, followed by a release command, and then the next is pressed.

In this scenario, the key press packets are each 3-bytes long while key release packets are each 1-byte long (see Example Packets). This gives an average packet length per command of 2-bytes – leading to an optimum baud rate (to ensure we do not exceed the 1000Hz USB protocol limitation) of 45000 baud.

As most other commands will generally result in equal, or longer, packets it is safe to assume that a maximum baud rate of 38400 (the nearest lower standard baud rate) will ensure the USB specification is never exceeded.

# Appendix F: Communications Protocol

Version 1.0

## Requirements

This protocol is required to include support for Mouse, Keyboard and Joystick emulation in a robust and high efficiency manner. In addition to this, the protocol should aim to be as understandable as possible and avoid complex behaviour which complicates implementations wherever possible.

For performance reasons this requires that the protocol be binary in nature, reducing (and in many cases removing) the need for packet parsing. In addition to this, attempts will be made to reduce the amount of data which will be transmitted over the UART connection to improve performance as much as possible.

## Design Decisions

There are two approaches to the protocol which we are able to take. The first is to attempt to design a protocol which is as faithful to the USB HID specification as possible - effectively causing the ATmega to act as a relay device, however while this will certainly minimize packet size to a large degree and faithfully allow emulation of any USB HID device it also has the distinct disadvantage of requiring the master implementation to handle the creation of all HID packets - a complex task which is prone to errors.

The simpler alternative is to rely on the included HID emulation libraries and instead declare a protocol which acts to perform RPC (remote procedure calls) on the ATmega chip. This, if well designed, could result in smaller packets for most common operations and would significantly simplify protocol implementations. The obvious disadvantage of this approach is that in order to emulate additional devices it would be necessary to extend the functionality of the ATmega's firmware as well as (possibly) adding additional op-codes to the protocol.

Version 1.0 of this protocol will adopt the second approach, attempting to implement a very specific RPC system built around USB HID emulation on the ATmega chip. Packets will consist of an op-code, packet length field and a number of 8-bit arguments to be passed to the corresponding functions. If needed, these 8-bit arguments can be combined to create 16-bit or 32-bit values where those are necessary.

## Protocol

### Packet

All protocol operations are wrapped in a packet structure similar to the following. Packets consist of a 3-bit **OP\_CODE** field, a 5-bit **ARG\_COUNT** field as well as a number of 8-bit arguments. There is a protocol imposed limit of 31 arguments, limiting the total packet size to 32-bytes.

Table UART Protocol Packet Structure

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Bit Index | **0x0** | **0x1** | **0x2** | **0x3** | **0x4** | **0x5** | **0x6** | **0x7** |
| **0x00** | OP\_CODE | | | ARG\_COUNT | | | | |
| **0x10** | ARG\_1 | | | | | | | |
| **0x20** | … | | | | | | | |
| **0x30** | ARG\_N | | | | | | | |

### OP Codes

There are a number of basic op codes which cover the spectrum of available functions which may be performed by the emulation layer. These codes define the way in which their received arguments are treated and allow future extensions to the protocol through the use of the **000** op code.

Table UART Protocol OP Codes

|  |  |
| --- | --- |
| OP\_CODE | Description |
| 0x0 000 | Custom Operation |
| 0x1 001 | Keyboard |
| 0x2 010 | Mouse |
| 0x3 011 | Joystick |
| 0x4 100 | Reserved For Future Expansion |
| 0x5 101 | Reserved For Future Expansion |
| 0x6 110 | Reserved For Future Expansion |
| 0x7 111 | Reserved For Future Expansion |

### Expected Arguments

Each op code expects a certain set of arguments to be provided, and their presence dictates the behaviour undertaken by the emulation layer when the op code is received. In all cases, transmission of a packet with **ARG\_COUNT=0** will be used to release all active keys or buttons.

#### Keyboard

The keyboard operation is used to trigger the emulation of KeyDown operations. It is important to note that unlike platform native emulation libraries like SendKey() or Win32 API calls it is not necessary to send a KeyUp message when performing USB emulation, rather the KeyUp state will be detected when a packet is sent without the key listed as depressed. This is an important distinction and one which will in many ways dictate the way this protocol is designed.

In addition to this, the ATmega32u4's keyboard emulation is limited to 6 keys + 4 modifiers at any one time, and due to the way that HID emulation is performed it is impossible to "trick" the operating system into believing that more than that number are depressed at any one time.

As a result of these restrictions, the Keyboard (**0x20** flag) OP\_CODE requires the following argument structure.

* **ARG\_1** MODIFIERS : **uint8**
* **ARG\_2..7** KEY : **uint8**

It is important to note that the protocol and implementation allow the transmission of partial packets - meaning that it is not necessary to send additional arguments for keys which are not in use. Therefore to send the 'A' key it is simply necessary to send 2 arguments.

##### Example Packets

The following are examples of packets for performing some basic operations. Numbers represent hexadecimal bytes sent to the emulation device

* **Press A** 22 00 04
  + 0x20 – Keyboard Operation + 0x02 – 2 byte packet length
  + 0x00 – No modifier keys
  + 0x04 – Key code for A key
* **Press Shift+A** 22 02 04
  + 0x20 – Keyboard Operation + 0x02 – 2 byte packet length
  + 0x02 – Shift modifier key
  + 0x04 – Key code for A key
* **Press Ctrl+Shift+A+B+C** 24 03 04 05 06
  + 0x20 – Keyboard Operation + 0x04 – 4 byte packet length
  + 0x03 – Ctrl and Shift modifier key codes
  + 0x04 – Key code for A key
  + 0x05 – Key code for B key
  + 0x06 – Key code for C key
* **Release All Keys** 20
  + 0x20 – Keyboard Operation

##### Known Issues

Because of the way emulation is implemented in the Teensy firmware, it is impossible to send the full **uint16\_t** key codes to the emulation layer. This means that it is not possible to emulate certain special keys like **VOLUME\_UP**, **MUTE** etc. at this time. In the future, if this functionality becomes available it may be possible to tweak this implementation to support sending the full key codes in which case modifiers and keys will need to be handled differently.

#### Mouse

The mouse operation type is used to emulate mouse button presses, movement and scrolling. As with the keyboard operation type, transmitting a Mouse packet with no arguments has the effect of releasing all pressed buttons.

Button presses are encoded into the first argument using a set of flags, namely the following. The button flags are OR-ed together to give the resulting button code.

* **Left** 0x1
* **Right** 0x2
* **Middle** 0x4
* **ARG\_1** BUTTONS : **uint8** flags
* **ARG\_2** DELTA\_X : **uint8**
* **ARG\_3** DELTA\_Y : **uint8**
* **ARG\_4** DELTA\_SCROLL : **uint8**

It is important to note that it is possible to send "partial" packets, in which case the subsequent values will be assigned a default value of 0. This means that a mouse button press emulation doesn't need to send the DELTA\_X, DELTA\_Y or DELTA\_SCROLL components. Similarly, a Y movement doesn't need to send the DELTA\_SCROLL component.

##### Example Packets

* **LMB Down** 41 01
* **Right 8px** 42 00 08
* **Scroll Up 2 Lines** 44 00 00 00 02
* **Reset Buttons** 40

##### Known Limitations

Due to the USB HID specification not supporting Mouse Button 4 or 5 (used commonly to provide Forward/Backward navigation) it is not possible to emulate these. In addition to this, the HID specification provides no way to move the mouse to an absolute position on the display (given a set of X, Y coordinates). This behaviour can be emulated by moving the mouse to the bottom left corner (repeated -127, -127 movements) followed by movements to the desired location. The number of movements required to move the mouse to (0,0) will depend on the target display's resolution.

#### Joystick

The joystick emulation layer is slightly more complex than that of the mouse or keyboard - as it is necessary to pack relatively more information into the packet than would otherwise be necessary. The ATmega32u4 is capable of emulating a joystick with 32 buttons, 6 axes and a single 8-way hat switch. In order to provide full accuracy (10-bit axis reporting) it is necessary to "pack" sets of 3 axes together such that one 4-byte integer contains axis information for 3 axes.

Packing is achieved by applying the following algorithm. Take note that 2-bits are lost for each set of 3 packed axes, resulting in a packed efficiency of 93.75%, compared to an efficiency of 62.5% if packing is not used.

int32\_t pack(int16\_t axis1, int16\_t axis2, int16\_t axis3) {

return (((axis1 << 10) | axis2) << 10) | axis3;

}

In addition to this, the hat switch is handled differently to the standard Arduino implementation to allow its data to be contained within a single 8-bit argument. The special value 0xff is used to represent center, while all other values are multiplied by 45 to give the number of degrees from north.

The resulting packet is in the form

* **ARG\_1..4** BUTTONS : **int32\_t**
* **ARG\_5..8** pack(X, Y, Z) : **int32\_t**
* **ARG\_9..12** pack(rZ, sL, sR) : **int32\_t**
* **ARG\_13** HAT : **int8\_t**

It is important to note that as with all other op-codes it is possible to send empty packets, however due to the way in which axes are handled this is not recommended under any circumstances as strange values will be reported. In future, the upper bit of a pack may be used to indicate that it is a valid value and should be updated, however that is currently beyond the scope of this implementation.

# Appendix G: C-Library API

The C API for Isotope comes bundled with an *isotope.h* header file which clearly documents all methods available for use from within your applications. The following is a quick reference sheet describing the different API methods available to you through the API.

*// Opens a new Isotope device connection*

**int** isotope\_open(**const char\*** device);

*// Closes an Isotope device connection*

**char** isotope\_close(**int** isotope);

*// Formats and sends a mouse command to the specified Isotope*

**char** isotope\_mouse(**int** isotope, **char** buttons, **char** deltaX, **char** deltaY, **char** deltaScroll);

*// Formats and sends a keyboard command to the specified Isotope*

**char** isotope\_keyboard(**int** isotope, **char** modifiers, **const char** keys[], **char** keys\_count);

*// Formats and sends a joystick command to the specified Isotope*

**char** isotope\_joystick(**int** isotope, **int** buttons, **short** x, **short** y, **short** z, **short** rz, **short** sliderLeft, **short** sliderRight, **char** hat);

*// Sends the specified packet to the specified Isotope*

**char** isotope\_write(**int** isotope, **const** **char\*** packet, **char** length);

# Appendix H: Source Code

All project source code, documentation and development has been undertaken on a private git repository available at <https://git.sierrasoftworks.com/stellenbosch/isotope>. For access to the repository, please contact Benjamin Pannell at [admin@sierrasoftworks.com](mailto:admin@sierrasoftworks.com) and reference this document.

Certain parts of the implementation, including the Node.js library, have been released under permissive open source licences and are publicly available on the following websites.

* Isotope for Node.js – <https://npmjs.org/package/libisotope>